Slide 1 of 46

Title Slide: Considerations for Analysis of ECLS-K Data

Slide 2 of 46

This module provides an overview of considerations for analysis of the ECLS-K data.

The ECLS-K data file contains a wealth of data from and about children, their parents, their teachers, and their education settings. This module provides information about how to use some of the data collected from multiple sources across the ECLS-K, as well as some considerations to keep in mind when conducting analyses with these data.

The first section of the module discusses the information available to identify language minority children and, more specifically, English language learners, along with the limitations of the ECLS-K data on this topic. The second section describes the information available in the dataset to identify children with disabilities and conduct research on the services they receive. Third, student mobility and how this can impact analyses are discussed. Fourth, the collection of information from teachers, in particular the structure of the teacher questionnaires, is described. Finally, the types of geographic information available and where that information is located are discussed.

There are two ways to access the information within this module. You can click on one of the links above, which will take you directly to that section of the module, or you can click the 'next' button to advance to the next slide in the module and view all of the analysis considerations presented in the order they are presented on this slide. At the end of each section of the module, you will be provided with two buttons, one that will return you to this objectives page where you may select another section of the training module, and one that will allow you to exit the module.

Slide 3 of 46

Information about children's home language was collected from both schools and parents. This information can be used to identify language minority children, or those whose native language is not English. Additionally, information collected from teachers and results from the ECLS-K assessments can be used to identify those language minority children who might be classified as English language learners, or ELLs for short. The following slides describe the available information about children's language in more detail.

Slide 4 of 46

In the primary language, or PLQ, section of the parent interview, parents were asked several questions related to languages spoken in the child's home. First they were asked whether any language other than English was spoken in the home. If so, they were asked whether English was also spoken. They were then asked to report the primary language spoken in the home. Finally, they were asked how often key household members used a non-English language when speaking with the child and,

relatedly, how often the child used a non-English language when speaking with those household members. Key household members were usually the child's mother and father figures or primary guardian or guardians.

Slide 5 of 46

There are two main sources of school-based information about home language and ELL status. First, school staff were asked to provide information from school records about children's primary home language. Second, children's classroom teachers were asked about languages used in their classroom and whether any of their students were language minorities or had limited English proficiency. Though the questions about language and proficiency were not asked about each study child specifically, if all of a teacher's students were language minorities or English language learners, it can be assumed the study child was as well. Teachers were asked to indicate whether each study child received English language instruction as part of an English-as-a-Second Language, or ESL, program and whether that instruction was received in the classroom or as part of a pull-out program. They also indicated whether the study child participated in a Title I-funded ESL or bilingual education program.

Slide 6 of 46

As discussed in the module titled, 'Data collected through the ECLS-K,' the direct child assessment included an English language proficiency screener called the Oral Language Development Scale, or OLDS. The screener was used in kindergarten and first grade to assess the basic English language skills of children identified by school records or teachers as having a non-English primary home language. Children who achieved the cut score or higher on the OLDS were administered the subject area assessments in English. Children who did not achieve at least the cut score on the OLDS were administered Spanish translations of the mathematics and psychomotor assessments if their home language was Spanish. They also had their height and weight measured. Children who did not achieve at least the cut score on the OLDS and whose home language was not Spanish had their height and weight measured. They were not assessed in the cognitive or psychomotor domains.

In each round of data collection during the kindergarten and first grade years, the OLDS was administered to children who were administered but had not passed it in the previous round. The OLDS was not used after the spring of first grade, because most children had passed it by that round. All children were assessed in English starting in third grade.

Slide 7 of 46

The OLDS consisted of three subtests of the *pre*LAS 2000: Simon Says, Art Show, and Let's Tell Stories. CTB/McGraw-Hill's *pre*LAS measures the English and Spanish language proficiency and pre-literacy skills of learners in early childhood. Together these subtests provide measures of children's listening comprehension, vocabulary, and ability to understand and produce language. Specifically, Simon Says measured

listening comprehension of basic English instructions. Art Show was a picture vocabulary assessment where children were asked to name pictures they were shown. The Art Show served as an assessment of a child's ability to produce language and measured the child's command of expressive language. The final part of the OLDS, Let's Tell Stories, was used to obtain a sample of a child's natural speech by asking a child to retell two stories read by the assessor. Whether a child achieved at least the cut score on the OLDS and was assessed in English can be used as an indicator of children's basic English proficiency. For children who were administered the OLDS, the data file includes a variable indicating a child's score on the OLDS for each round in which the child was administered it and a variable indicating in which round the child passed the OLDS.

Slide 8 of 46

Researchers interested in analyzing population subgroups such as language minority children are typically concerned with the size of the sample and whether it is large enough to support analyses. The determination of how many language minority children participated in the ECLS-K and would be included in analyses depends on how language minority children are identified. Identification depends on the research question being addressed. For example, researchers may be interested in looking at all language minority children, whether or not English is also used at home, a subset who do not speak English at home at all, or the subset of language minority children who are classified as English language learners.

The number of language minority children or English language learners also depends on the source from which information was obtained. For example, in the fall of kindergarten, about 2,850 children were identified through school records as having a non-English primary home language, while about 4,310 children had parents who said a non-English language was spoken in the home.

Looking at possible ways to identify ELLs, in the spring of kindergarten, about 1,680 children had teachers who said the child participated in an ESL program during the school year, while about 1,450 children did not pass the OLDS in the fall kindergarten data collection and about 960 did not pass in the spring kindergarten data collection.

Slide 9 of 46

Several issues related to identification of language minority children and English language learners are important to consider when conducting analyses. First, as just discussed, these populations can be identified in several ways using data from different sources, which may be inconsistent with one another. For example, a child who speaks a language other than English at home may not be identified as a language minority in school records if the child is proficient in English. Second, these populations can be identified based on initial status at kindergarten entry or status in the rounds of analytic interest. For example, the population of children identified as English language learners using scores on the OLDS or teacher reports of participation in ESL would be smaller in the spring of first grade than in the fall of kindergarten because the study children

acquired greater English language proficiency as the study progressed. Third, the ECLS-K includes large samples of language minority children, but the group is diverse. The majority of language minority children speak Spanish. Sample sizes for language minority children who speak a language other than Spanish may be too small to support detailed analysis.

Slide 10 of 46

Analysts should also keep in mind that the number of language minority children with complete assessment data increases across time. The assessment protocol resulted in children who did not achieve the OLDs cut score having missing or incomplete data in one or more of the earlier rounds of the study. This affects researchers' ability to measure achievement gains, because gains can only be calculated starting from the time point at which children passed the OLDS. Also, Spanish-speakers switched from Spanish to English math assessments once they passed the OLDS. This may have affected measurement of gain in mathematics knowledge and skills. Lastly, inclusion in later rounds of children who had not passed the OLDS at some earlier time point can affect analyses for certain race/ethnicity groups. Such children are likely to achieve lower scores than children whose native language was English or who passed the OLDS early in the study, so their concentration in certain race/ethnicity groups, for example Hispanics, can result in lower average scores for those groups in later rounds.

Researchers interested in examining the cognitive knowledge and skills of language minority children who do not have direct assessment data at all time points can consider using the Academic Rating Scale information reported by teachers as an alternative source of information. The Academic Rating Scale is described in the module titled, 'Data Collected through the ECLS-K.'

This concludes the section of the module on identifying language minority children and English language learners. To return to the objectives slide to make another selection, click the "Return to List" button. To exit the module completely, click the "Exit" button or you will automatically be advanced to the next slide within the module.

Slide 11 of 46

The ECLS-K collected a good deal of information about children with disabilities, including their specific disability and any services they received. This section of the module provides information about the multiple sources of data that can be used to identify children with disabilities and examine their educational experiences and outcomes, along with some considerations to keep in mind when using these sources.

Parents were asked many questions about any disabilities that their child may have had as well as receipt of services in the child health, or CHQ, section of the parent interview. A second source of information about children's disabilities and receipt of services were special education teachers and related service providers of study children who had an Individualized Education Program, or IEP, on record with the school. The third source of

information on child disability and receipt of services is administrative data, specifically the student record abstract and the field management system.

Some of the parent-reported and administrative data are available in the public-use file. However, much of the information on children with disabilities is only available in the restricted-use files, due to the sensitive nature of the data and the fact that some of the information is applicable to a relatively small group of the study children. A primary factor determining whether data are in the public-use file is the number of children with different characteristics for whom data are available, and the number can change across rounds of data collection. As a result, some data that are publicly available for one round are restricted in other rounds. Data collected from special education teachers and related service providers are only available in restricted-use format.

The next few slides discuss these three sources of information about disability and services in more detail.

Slide 12 of 46

There are several considerations to keep in mind when identifying ECLS-K children with disabilities. As just discussed, there are multiple sources of data that can be used to identify these children, including parents, teachers, and administrative data.

Further, there may be differences in the data collected from the various sources because of misreport or differences in how the questions were asked. For example, parents were asked whether their child ever had any disability, while school staff were asked only about disabilities for which children were receiving services. Since not all disabilities require that a child receive educational services at school, the number of children with a disability based on parent report varies from the number of children with a disability based on school report.

The sources of data that are used to identify children with disabilities should be determined by the research question being asked.

Slide 13 of 46

For the most part, measures of disability status and special education were repeated at each round, with two exceptions. First, no questions about disability or participation in special education were asked in the spring kindergarten or fall first-grade data collections. Two data collections occurred in these grades, and the questions about disability and special education were asked in the other round of that school year, that is in the fall of kindergarten and the spring of first grade. Second, no information about disability or special education was collected from school records in the eighth-grade data collection, because the student record abstract was not used in that round.

Slide 14 of 46

Questions in the parent interviews focused on any issues children had with attention, learning, activity level, communication, hearing, vision, and behavior. For each of these issues, parents were asked how their child compared to other children of the same age, for example with respect to his or her level of activity, whether the parent had any concerns related to the issue, whether an evaluation was sought from a professional, and whether the professional diagnosed a specific problem related to the issue. If a diagnosis had been received, the parent was asked what the diagnosis was, when the diagnosis was made, and whether the child had ever received services for the diagnosed condition.

The base year of the study included the most detailed questions related to disability. Both retrospective and current information about children's experiences with disability and services, from birth up through kindergarten, were collected. In later rounds of data collection, the questions were asked as updates; that is, parents were asked about any new diagnoses that the child received since the time of the last interview.

Slide 15 of 46

The special education teachers and related service providers of study children who had an IEP on record with the school were asked to provide information about themselves, such as their age, sex, education, and work experience, as well as information about the child.

Information that was collected specific to the study child included: the child's disability category, both the primary disability and any disability for which services were being received; the child's IEP goals for the school year; the extent and type of services provided for the year; the child's primary classroom placement; the teaching practices, methods, and materials used with the child; any assistive technologies used by the child; general education goals, academic expectations, and participation in assessments; any collaboration or communication the special education teacher had with the child's general education teacher; the special education teacher's frequency of communicating with the child's parents; and formal individual evaluations, such as psychological, speech/language, and learning/educational evaluations.

Note that while some general topics have been covered in all grades, the questions asked were not necessarily the same in each collection.

Slide 16 of 46

The data file includes two variables for each round in which the special education teacher component was included that indicate whether a study child has data from a special education teacher. When information about the teacher, which was collected in part A of the questionnaire, is available, the value of DxSETQA is 1. When information about the study child, which was collected in part B of the questionnaire, is available, the value of ExSETQB is 1. Information collected from the special education teacher in the base year is included in a supplementary restricted-use file that must be merged to

the main data file by the child ID. Special education teacher questionnaire data collected in subsequent rounds are included in the main restricted-use files that contain other data collected for the child.

Slide 17 of 46

The field management system, or FMS, was a database used by field staff to manage data collection. Before data collection for each round began, field supervisors asked school staff if study children were receiving special education services. This information was recorded in the FMS. Additionally, school records were consulted to determine whether the study children had an IEP on record with the school and what disabilities the children had. This information was recorded in the student record abstract, or SRA.

Slide 18 of 46

This slide identifies several variables that are available on the data file that provide some indication of whether a study child had any disability. A composite variable based on parent-reported information was created for each year of the study. Information about the derivation of this composite can be found in the user's manual and in an errata that was issued correcting an error in the composite after the release of the third-grade data. The corrected disability composites for kindergarten through third grade, which begin with an R for "revised," are only available from the errata; they are not included in the Kindergarten Through Eighth-grade Full Sample Public-use Data File.

The second set of variables includes those indicating whether a child has special education teacher data. If a child has special education teacher data, this is an indication that the child receives special education services. However, there is some teacher nonresponse to these questionnaires, so these variables should not be used as a definitive indicator of receipt of special education. A better source of information about receipt of special education services is the administrative data. However, those records may also have some missing data, so it may be best to use multiple sources.

Earlier it was noted that there are differences in the data collected from the various sources of data because of misreport or differences in the questions asked. In the base year, about 2,135 children had a disability according to parent reports, about 1,020 children had an IEP on record with school according to the student record abstract, about 730 children had data for special education teacher Part A, and about 700 children had data for special education teacher Part B to about 700 children had data from special education teachers about the children's disabilities and services received in school.

Slide 19 of 46

There are also several variables available on the data file that indicate the specific disabilities children have.

There is one variable for each of the specific issues asked about in the parent interview indicating whether a diagnosis was received for specific problems. For example, the variable P1DIAGNO indicates whether the child was diagnosed as having a learning problem by a professional prior to the fall kindergarten parent interview. In the fall kindergarten parent interview, questions were asked about receipt of services for specific types of problems, and there is one variable related to each of these problems.

Variables derived from questions asked in the special education teacher questionnaire include those indicating the child's primary disability identified in the IEP and all disabilities for which a child received services.

Lastly, there are variables derived from the student record abstract in each grade except for eighth indicating all the disabilities that were identified in the child's most current IEP.

Though the ECLS-K sample is large enough to allow for analyses of children with disabilities or who receive special education services as a broad group, researchers interested in children with specific types of disabilities, for example autism or ADHD, may find that the sample sizes for these more specific groups are too small to support analyses. Tables presenting sample sizes by specific type of disability can be viewed by clicking on the underlined screen text on this slide.

This concludes the section of the module on identifying children with disabilities. To return to the objectives slide to make another selection, click the "Return to List" button. To exit the module completely, click the "Exit" button or you will automatically be advanced to the next slide within the module.

Slide 20 of 46

Sample mobility and the ability to locate and follow sample members across time are issues in any longitudinal study. This section of the module discusses these issues specifically in the context of the ECLS-K. The term "mobility" as discussed here refers to a change in teacher or school, not residential mobility. As mentioned in other ECLS-K modules, due to the cost of following children who transferred from the school they attended in kindergarten, subsampling was used to determine which of these children would be followed for continued participation in the study.

Analysts investigating the relationship of school or teacher characteristics to child outcomes over time should consider how mobility may impact their findings. Analysts should also consider how the reduction in sample over time may affect the sample size available for analyses, particularly if the analyses focus on subgroups of movers.

Slide 21 of 46

As can be seen in this table showing the cumulative percentage of children in each data collection round who had changed schools after the fall of kindergarten, by eighth grade most children- 84 percent of the sample and 94 percent of the population- had changed

schools at least once. Almost two-thirds of the population of children had changed schools at least once by the time of the fifth-grade data collection.

Mobility is an important consideration because some children are more likely to change schools than other children. Race/ethnicity is an example of a characteristic associated with mobility. As can be seen here, in every spring data collection round except for third grade, a higher percentage of Black children than children of other races/ethnicities had changed schools at least once.

Slide 22 of 46

Any analysis looking at children's growth by school characteristics needs to account for student mobility, as do analyses of children's status in either fall or spring within a single grade that use information on school characteristics collected in the other round of data collection for that grade. Additionally, some analysts may want to use school information collected in a prior round to look at children's status in a later round when school information from the later round of interest is not available. In all of these situations, it is recommended that all children who changed schools between the first and last time points included in the analyses be identified, and then the analyst should do one of two things: either drop those cases so that the analysis is run only with children who remained in the same school across the time points included in the analysis or drop these cases, retaining only those whose new school had the same characteristics of interest as the previous school. For example, in an analysis examining the relationship between school sector and academic achievement, the analyst could either drop all cases in which children changed schools or drop only those cases for which a change in school resulted in a change in school sector. In the latter example, children who moved from one private school to another or from one public school to another would be retained, while children who changed from a public to a private school or from a private to a public school would be dropped.

Slide 23 of 46

Similarly, any analysis looking at children's growth from fall to spring within a single grade by teacher or classroom characteristics needs to account for student mobility, as do analyses of children's status in either fall or spring within a single grade that use information on teacher or classroom characteristics collected in the other round of data collection for that grade. It is recommended that all children who changed teachers from fall to spring be identified when conducting such analyses. Again, the analyst could then either drop those cases so that the analysis is run only with children who had the same teacher or were in the same classroom for the entire year or drop these cases, retaining only those whose new teacher had the same characteristics of interest as the previous teacher.

Slide 24 of 46

The data file includes a series of variables that can be used to identify children who changed schools. For example, variable FKCHGSCH indicates whether a child changed schools between rounds 1 and 2, or the fall and spring of kindergarten.

Slide 25 of 46

Similar variables identifying a change in school across other rounds of data collection are also provided on the data file. These later-round variables all end with the characters “SCHG.” The first four characters of the variable name indicate the rounds between which the change occurred. For example, R3R2 indicates the change being noted occurred between round 2, the spring of kindergarten, and round 3, the fall of first grade.

There are also first-grade variables indicating school change between rounds in first grade and school change between the spring kindergarten collection and each first grade collection. For later rounds, the school change variables indicate whether a child changed schools between a given round and the round immediately preceding it. These variables from the later rounds also indicate whether the change involved a change in school sector. That is, these variables indicate whether a child transferred from a public school to another public school, from a public school to a private school, from a private school to another private school, or from a private school to a public school.

Slide 26 of 46

The data file also includes two variables that can be used to identify children who changed teachers. Variable FKCHGTCH indicates whether a child changed teachers between rounds 1 and 2, and R4R2TCHG indicates whether a child changed teachers between rounds 2 and 4. There are no variables indicating a change in teachers that include round 3, the fall first grade data collection. This is because no data were collected from teachers in round 3. Also, there are no teacher change variables that include rounds 5, 6, or 7, which were all spring data collections, because there were no fall collections in those grades and most children changed teachers from grade to grade. However, in the spring data collections for every grade teachers were asked whether the study child had been in his or her classroom since the beginning of the school year. Information from that question can be used to identify children who changed teachers during the school year. Note that any child who changed schools also changed teachers.

The change variables were created by comparing a child’s school or teacher ID from one round to another. When the ID does not match from one round to the next, this indicates that the child changed schools or teachers. Users can create these same types of composite variables themselves, because round-specific ID variables are included in the data file. These composites were included in the file to facilitate analysis of mobility.

This concludes the section of the module on student mobility. To return to the objectives slide to make another selection, click the “Return to List” button. To exit the module completely, click the “Exit” button or you will automatically be advanced to the next slide within the module.

Slide 27 of 46

The study used the same general procedures to collect information from teachers in the kindergarten, first-grade, and third-grade rounds of data collection since children generally stayed in the same classroom all day and had a primary teacher. There was a shift in study procedures beginning in fifth grade, when more children were switching teachers during the school day for instruction in different subjects. This module discusses the structure of the teacher questionnaires and data collection procedures, because they have implications for how the data are used and interpreted.

In the kindergarten, first-grade, and third-grade data collections, the sampled children’s primary or homeroom teachers reported information about themselves and their teaching practices. This primary teacher reported on instructional practices in all core subject areas, even if the child was not in the homeroom teacher’s classroom for instruction in a particular subject. For example, if a child left the primary teacher’s classroom for a given time period every day to be taught math by another teacher, information was still obtained about the primary teacher’s math instruction. This data collection strategy was used for two main reasons. First, changing teachers for instruction is fairly uncommon in the early grades. Second, due to clustering within classrooms, there were likely to be other study children to whom the primary teacher did provide subject area instruction, even if some study children left his or her classroom. As part of the child-level teacher questionnaire, teachers were asked if they were the child's primary teacher for reading, mathematics, science, and social studies, so this information can be used to identify children for whom the information about instruction in the different subject areas was not obtained from the teacher who provided such instruction to that child.

Slide 28 of 46

By fifth grade, a substantial number of children were receiving instruction in different subject areas from different teachers, so the data collection model was altered. Separate questionnaires were given to the study children’s teachers in different subjects to collect data on the instructional practices that children experienced in those specific teachers’ classrooms. All study children’s reading/language arts teachers were surveyed. Note that in the eighth-grade collection, the study referred to this teacher as the English teacher. In fifth grade, children were randomly assigned to have either their math or science teacher complete a child-level questionnaire. Thus, for a randomly sampled half of the children, the child’s mathematics teacher was surveyed. For the other randomly sampled half of the children, the child’s science teacher was surveyed. In eighth grade, for each child a questionnaire was given to the teacher for the same subject area that the child was sampled for in fifth grade. This procedure was used so that children would have two rounds of math or two rounds of science teacher-provided

child-level data. In any case in which the reading teacher also taught math or science, the teacher was given both the reading and the math or science questionnaires to fill out for that child.

Slide 29 of 46

In all rounds of data collection, the teacher questionnaire consisted of multiple parts. In the kindergarten, first-grade, and third-grade collections, each child's primary teacher was asked to complete three separate questionnaires. There was a teacher-/classroom-level questionnaire with questions about classroom characteristics and the teacher's instructional and grading practices. This is referred to as teacher questionnaire A, or TQA. There was also a teacher-level questionnaire with questions about the teacher's background and experience, views on teaching, job satisfaction, and views on the school climate and environment. This is referred to as teacher questionnaire B, or TQB. Teachers completed TQA and TQB only once in the rounds in which these questionnaires were fielded. A third questionnaire, referred to as teacher questionnaire C, or TQC, included questions specifically about the study child. For example, the indirect assessment questions about children's socioemotional functioning and the Academic Rating Scale were in teacher questionnaire C. Teachers were asked to complete one TQC for each study child in their classroom. With this questionnaire structure in the early grades, information was obtained about what and how the teacher was teaching, but if a study child was being instructed in one or more subjects by a teacher other than the primary teacher, the information collected was not necessarily describing the teaching practices experienced by the study child for those subjects.

Slide 30 of 46

In the fifth- and eighth-grade data collections, the questions in TQA and TQB were combined into one teacher-/classroom-level questionnaire. This questionnaire was given to all reading, math, and science teachers with a sampled child in his or her classroom. A separate child-level questionnaire, the TQC, was again fielded, but the TQC was customized according to the subject area the teacher taught. The TQC for reading had questions specifically about the child's experiences related to reading, including the child's reading achievement, the reading teacher's instructional practices, and the characteristics of the classroom in which the child received reading instruction. Similarly, the TQC for mathematics had questions specifically about the child's experiences related to math, and the TQC for science had questions specifically about the child's experiences related to science. With this questionnaire structure in the later grades, information obtained from teachers described the teaching practices experienced by the study child for the TQC focal subject.

Slide 31 of 46

The variable naming conventions used for the variables derived from teacher data in kindergarten, first grade, and third grade correspond to the structure of the questionnaires as they were fielded. Variables that begin with an **A** or **B** prefix are based on information from the teacher-/classroom-level questionnaires. More

specifically, variables beginning with **A** are derived from TQA and variables beginning with **B** are derived from TQB. Variables that begin with a **T** prefix are based on information from the child-level questionnaire that is specific to the sample child. The reason the TQC variables start with a **T** instead of a **C** is that the child assessment variables begin with a **C** prefix.

Slide 32 of 46

The variable naming conventions used for the variables derived from teacher data in fifth and eighth grade also correspond to the structure of the questionnaires as they were fielded. Therefore, the naming conventions are somewhat different in these two rounds compared to prior rounds. All variables that begin with a **J** prefix are based on information collected from the teacher-/classroom-level questionnaire. As discussed in the module titled 'Getting Started with the Data,' the second character in a variable name indicates the round of data collection, in this case "6" for round 6, or the fifth-grade round, and a "7" for round 7, or the eighth-grade round. The third character of a variable beginning with **J** is a number associated with the subject area taught by the teacher who completed the questionnaire. A "1" is used for information reported by a reading/language arts teacher, and a "2" is used for information reported by a mathematics or a science teacher.

Slide 33 of 46

Variables that begin with a **G, M, or N** prefix are based on information from the child-level questionnaire that is specific to the sample child, with each prefix indicating the focal academic subject for the questionnaire. **G** is used for reading, **M** is used for mathematics, and **N** is used for science.

Slide 34 of 46

Though the presence of data in either the M or N variables is an indicator of whether it was the study child's mathematics or science teacher who was sampled to participate in the study, the data file does include a variable for each round of data collection indicating which teacher was sampled. **F6MTHSCI** is the variable for fifth grade and **F7MTHSCI** is the variable for eighth grade. As noted earlier, for any given child, the subject sampled for fifth grade was also sampled for eighth grade, so **F6MTHSCI** and **F7MTHSCI** should identify the same subject for children who have these data in both rounds.

Slide 35 of 46

When running analyses using information from teachers, there are some issues related to the structure of the teacher questionnaires and the procedures used to collect data that should be considered. Many questions asked of teachers were repeated over time; however, some were asked at the teacher level in the earlier grades and then were asked at the child level in the later grades. As a result, in the early grades, teachers may have reported about classroom practices that were not experienced by every study child in the class. In the fifth and eighth grades, child-level mathematics and science

teacher data are not available for all sample children. About half of the children have mathematics data and about half have science data. The data file includes analytic weights to be used with teacher data in these later rounds that, when applied, adjust for the subsampling that was used. Weighted data are nationally representative even though math and science information is only available for half the sample. However, the smaller sample sizes may not support analyses related to mathematics and science conducted for smaller subgroups within the study population, for example children with disabilities.

This concludes the section of the module about teacher questionnaire structure and data collection. To return to the objectives slide to make another selection, click the "Return to List" button. To exit the module completely, click the "Exit" button or you will automatically be advanced to the next slide within the module.

Slide 36 of 46

Many researchers are interested in geocode information in order to conduct analyses relating where children live to children's characteristics, experiences, and outcomes. The ECLS-K data files include various indicators of geographic location, though some information is only available in restricted-use format. This section of the module discusses what geographic identifiers are available and in what files they can be found.

There are only two variables with geographic area information available in the public-use file. One indicates Census region, which uses categories corresponding to four areas of the United States: Northeast, Midwest, South, and West. The other is a locale variable, indicating whether the area in which a child lives is a city, suburb, town, or rural area. More specific geographic data are available on the restricted-use data files.

Slide 37 of 46

These are the various geographic identifiers available for children's schools and homes. The Federal information processing standards, or FIPS, codes can be used to identify the state and county in which the school and home are located. Standard ZIP codes are also provided, along with another identifier called a ZCTA, or ZIP code tabulation area. ZCTAs are areas created from Census blocks that are virtually identical to regular ZIP codes but include ZIP codes associated with entities that are too small to have data released for them, such as a ZIP code consisting of only one post office with P.O. boxes or a large apartment building with its own ZIP code that is located within a larger ZIP code. Variables identifying Census tract are available for homes and schools. Lastly, latitude and longitude are available for schools.

Slide 38 of 46

Restricted-use geocode data are available in two locations: a supplementary set of geocode data files that was created for the kindergarten through third-grade rounds of data collection, and the regular cross-sectional restricted-use files for first through eighth grade that are discussed in the module titled, 'Getting Started with the Data.'

The geocode data CD, with NCES number 2004-116, has data files that include all the geographic identifiers discussed on the previous slide for the kindergarten, first-grade, and third-grade rounds. A user's manual describing the data and how to use them is also provided on the CD.

Slide 39 of 46

The geocode data files also include more than 600 variables derived from 2000 Census data that are related to the demographic characteristics for each ZCTA and tract corresponding to schools and children in the ECLS-K sample. Examples of variables included in the file include: the median average income of residents, the percent of residents in different race/ethnicity and age groups, the percent of residents who own or rent their homes, the percent of households that are headed by females, and the percent that are military households.

Slide 40 of 46

Other Census Bureau data can be obtained from the Census website and linked to the ECLS-K data by ZCTA or tract. Examples of such data include the proximity of a child's home or school to social service agencies, grocery stores, industries, parks, and recreational facilities in the area.

Slide 41 of 46

The geocode data CD comes with eight ASCII data files that include different pieces of the available information. The file **school.dat** contains tract and ZCTA identifiers for the schools that participated in at least one round of data collection from kindergarten to third grade. The files **child1.dat, child2.dat, child3.dat, child4.dat, and child5.dat** contain tract and ZCTA identifiers for the ECLS-K children's homes in rounds 1, 2, 3, 4, and 5, respectively. There is one file for each round because children's home addresses can differ from round to round. The file **tracts.dat** contains 635 Census-derived variables for all tracts in the United States. The file **zctas.dat** contains 635 Census-derived variables for all ZCTAs in the United States.

Slide 42 of 46

The geocoded files can be linked to the main ECLS-K data files using the school ID and child ID variables. Additional Census data can be linked to the geocode data files using the ZCTA and tract variables. Note that the 11-digit tract variable includes the state FIPS Code (the first 2 characters), the county FIPS Code (the next 3 characters), and the Census tract code (the last 6 characters).

Slide 43 of 46

Beginning with the first-grade restricted-use file, the cross-sectional restricted-use data file includes the state FIPS code, county FIPS code, and ZIP codes for the schools, as well as ZIP codes for the children's homes. For eighth grade, the restricted-use data file also includes ZCTA and tract data for schools and homes, and longitude and latitude

data for schools. There are no ZCTA, tract, or longitude and latitude data available for the fifth-grade round of data collection, although these identifiers can be obtained from other rounds for children who have not moved homes or changed schools.

Slide 44 of 46

This chart summarizes the more detailed geographic identifiers that are available for each round of data collection and indicates the file where such information can be found. To summarize, geographic codes can be found on the geocode CD for kindergarten to third grade, along with many community demographic variables derived from Census data. Geocode data are also available on the cross-sectional restricted-use files starting in first grade. Census data can be linked to the ECLS-K data using these geographic area codes. No ZCTA, tract, or longitude and latitude data are available for the fifth-grade round of data collection.

This concludes the section of the module on geographic area data. To return to the objectives slide to make another selection, click the “Return to List” button. To exit the module completely, click the “Exit” button or you will automatically be advanced to the next slide within the module.

Slide 45 of 46

This module has provided you with an overview of considerations for analysis of the ECLS-K data. Specifically, analysis considerations for identifying language minority children, English language learners, and children with disabilities were described. Information was also presented about special education data, student mobility, teacher reports and the corresponding teacher questionnaires, and the availability and location of geographic data.

Slide 46 of 46

Important resources that have been provided throughout the module are summarized here for your reference.

You may now exit the module.